

Weekly Report

Lu Junhua

2016 年 7 月 11 日

Done

- Designed a model illustration picture as 图 1.
- Revised the model part of our part and got it revised by Prof. Peng.
- Rewrote a writing plan of VIS paper according to the reviews, as follows.
- Discussed with Dr. Xie on the design plan of VIS paper, thought about new visual encodings, and the reason of adopting this design.
- Study the first four chapters of *Web Design with JavaScript and the DOM*.

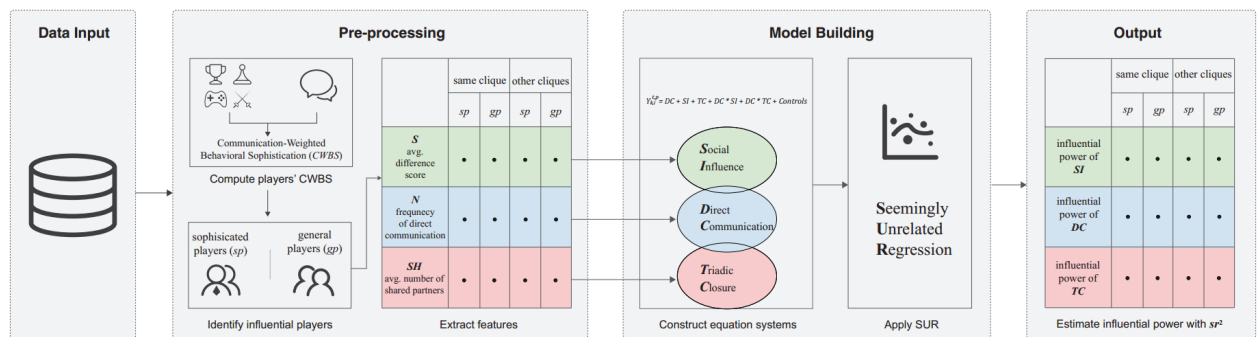


图 1: 模型示意图

BeXplorer: Writing Review

Easy

- Description of visual elements:** Where is that information the colored bars? It is really confusing - as I can not see these differences. (Primary)

Suggestions: colored bars 改成 influence bar 可能就看得懂了?
- Description of visual elements:** Color of the nodes shows "aggregated influence" on the corresponding flow. I'm not sure what that means, and there is no scale. It only appears to be relevant for the middle tier of the tree as well. The tree mixes the types of its elements with materials rather than players on the lower tier (I would think a mechanism that switches between player centric or material centric views would be easier). The use of the orange line to indicate selection is non-standard. It is not clear how the flows can be easily matched back to the selections. What would we make of flows from selections at different levels? (Secondary)

Suggestions: 我们对设计进行了修改. 对于"aggregated influence", 如果还要用的话可以详细解释下, 这个 aggregate 是指整个 group 受到的影响力的和.
- Description of visual elements:** Figure 1 also makes it look as if these are a reflection of a particular point in time, but after watching the video, I'm not sure that is the case. (Secondary)

Suggestions: 我们的视频中已经标出了这一点(select a time frame), 这个问题我认为可以直接 argue.
- Description of visual elements:** The mechanics of the distribution ring are similar obscure. I do think it would help to use figures in which the weren't basically the same across all of the populations so we could see some variation. I'm also not sure how to read the parallel coordinates because it isn't clear what the scales are. (Secondary)

Suggestions: (a) 已经设计做出修改, 并且会选择差异更大的图来展示. (b) PCP 加上标注, 说清楚我们是把数据归一化了.
- Descriptions:** 缺少 labels. | The descriptions of the visual representations (in Section 5) are confusing and hard to follow. For example, in Figure 2, the authors illustrate the distribution ring visualization but it is very difficult to understand what each layer represents. I would suggest taking on a specific attribute as an example and labeling each portion of Figure 2 with as much detail as possible. In fact, the lack of labeling in the supplemental video, the figures, and on the visualization dashboard itself is troubling. Where do the labels (B,M,T,G) come from? Are they arbitrary? What do they mean in this context? In the paper, the authors simply state that there are 4 cliques and they are B, M, T, and G. If these four labels are the most significant icon that act as "anchors" for which each panel is related to the other, the icons (or alpha-numerical symbols) should have inherent meaning that is easy to understand. (Secondary | External 2)

Suggestions: 事关 distribution ring(现在设计已经改了, 但是的确这个标注非常费解, 这个要重写标注.) | 视频中的 labels 也要注意, 不要与系统中自带的混淆, 要多加几个 labels. | BMTG 怎么来的, 这个文中已经提到了, argue.
- Minor issues:** The introduction refers to the detail-level view as the "low-level" view. Revise this to be consistent throughout the paper. | The equation font appears to be smaller than required by the VIS template and should be revised. | PCPs shows.. / Three PCPs...is used.

(External 1)

Suggestions: 按照这个要求来就行. 但是方程的字体修改可能导致公式完全变丑.

Medium

1. 图片位置不好(Primary)

Suggestions: 这个有待商榷. 后面若能调整当然是尽量放好一点.

2. **Description of visual elements:** The influence bars are a straightforward enough visualization, but it is not clear to me how they determine the influence based on these isolated components of social interaction. (Secondary)

Suggestion: Sec. 4 新增的模型示意图应该已经能帮助读者理解了. Influence bar 和 influential power 没有直接联系, 这在 Sec 4 末尾已经提到. 但是可能需要在可视设计那边重新提一下, 指向 Sec 4 的这几个部分.

Hard

1. **Model:** It would be much more efficient if you could explain first in high level what "influence-susceptible model" is and how it contributes to your purposes. The description you provide was not sufficient for me. (Primary)

Suggestion: high-level 解释一下我们的模型及对我们所做的工作的作用

2. **Design goals:** the design goals are interesting but very specific to the game you worked with. How much of what you describe is transferable to other domains? Would it be possible to find here a task taxonomy to back up your description? (Primary)

Suggestion: design goals 能否运用到别的领域, 这个可以 argue

3. **Design goals:** I like that you tied the tasks to the different views, but subdivision is unclear. (Primary)

Suggestions: 个人的理解是各个 Views 对 Design goals 的分配不够好. High-level view 的细节太多了算不上 high-level view.

4. **User Feedback:** the user feedback is weak in my opinion. a reflection on design decisions would have been more effective, and I also think that for the current case, a scalability or applicability to other data would have been more insightful. | though the tool and the model seem to be directed at evaluating this particular game, and it is not clear how either would be easily generalized. | However, the evaluative pieces of the manuscript appear to only use one dataset, which may significantly limit understanding of the generalizability of the techniques introduced. (Primary | Secondary | External 1)

Suggestions: (a)user feedback 肯定是要重新去做的. (b)对于 design decision 的反思, 如果一定要写, 可以结合新的 feedback 写. (c)对于不同数据的扩展性/适用性/普适性等很多人提到了, 我们 argue, 按照之前的计划, 我们在 introduction 里面会提一下, 再在 discussion 中细讲我们其性质. 包括的应该有 CWBS 的普适性, 模型的普适性.

5. **Model:** As a result, I found sections of this work (like section 4) almost impenetrable. I would encourage the authors to consider the audience, and think about making the work accessible to readers that are not, for example, familiar with triadic closure. *For that matter*, just knowing more about the game itself (e.g., what are the players consuming?) might help to make concepts

more concrete (and explain why there is influence/ communication between cliques).
(Secondary)

Suggestions: (a)三个基本概念解释一下于 Sec. 4. (b) 他们买的东西之前有写到, 只是因为篇幅原因被删了. 但是我认为这样加上也并不会太改善他们对于这个模型的理解 (针对 *for that matter* 这个短语). 希望(a)已经能够满足需求.

6. **Case study:** it is curious that the first case study is only peripherally related to the original questions. (Secondary)

Suggestions: 第一个 case 和最初的 design goals 没有紧密的联系. 这一点我认为是 case 写的时候没有顺带标出解决了 Q 几 Q 几导致的, 应该来说这个 case 和 design goals 是紧密的, 除了没有与商品这个 Q 相关.

7. **Design goals:** I'm also not sure that I see a big distinction between Q1 and Q2. The questions themselves seem different, but the explanatory text blurs the lines between them. | If the explanations regarding the visual design were more clear (Secondary | External 2)

Suggestions: Q1 和 Q2 两者差距的解释不够清晰, 这需要语言的重新修改, 比如删除 Q1 那个 multiplex 这个词.

8. **User feedback:** The use cases are helpful, but I'm not sure how much the interviews with the analysts add to the work. | the authors provide evidence that domain experts can, in fact, make use of the series of visualizations to successfully complete targeted tasks. / I am skeptical of the user feedback descriptions (Section 6.3). Without full immersion and a detailed (and controlled) set of tasks, how can we properly vet the domain expert's opinion? In order words, how do we know that the three analysts are validating the case study descriptions and not the visualization or interface? Also, there is a lack of description in the interview itself. The authors refer to it as a semi-structured interview and interface walk-through. It seems like this approach would naturally elicit biased, positive responses. (Secondary | External 2)

Suggestions: 改设计重新做 interview. 这一块要写的更详细真实一些, 具体要等后面了. 按照 Ext2 的疑惑可以改进后面的 interview, 比如设计好一些任务让他们去完成.

9. **Description:** In summary, the paper is well-written, but does not properly describe the relationships between the coordinated visualizations. If Section 5 was as clear as its predecessors, the paper would be acceptable for publication. (External 2)

Suggestions: 尽管这里用了 in summary, 我并不能看到他前文有提到 relationship 的词, 这个可能需要在整体描述中蕴含这样的一层含义, 或者显示的说清楚从高层到低层, 从大局到细节的这样一个关系.

To do

- Continue revising the paper according to the plan.
- Proceed to learn JavaScript according to this book.

论文

- PVIS15 *VisMOOC: Visualizing Video Clickstream Data from Massive Open Online Courses* 本文是一个prototype, 其网站上的内容结合了本文以及下面这篇文章的一些元素和视图. 这篇文章主要做的还是对一些事件的展示, 这里的事件包括了play pause seek stalled ratechange和error. 其着重展示的有seek graph, 表现了用户往前找还是往后找的一些信息. 这种类似于调用关系的图其实是很复杂的, 除了颜色区分外, 作者还用了一种称为**adaptive transparency**的技术, 使得这种交错变得没那么严重, 并且如果存在明显的pattern的话, 将是很容易找到的.
- TVCG *PeakVizor: Visual Analytics of Peaks in Video Clickstreams from Massive Open Online Courses* 文章初衷是因为之前开发的系统VisMOOC在被专家使用时, 专家对那些数据中的peak非常的感兴趣. 文章用的数据主要是玩家与课程视频交互的一些数据比如暂停、找寻、播放等等, 并且结合了课程本身的信息, 以及用户的一些信息. 地理位置信息同样的也被在地图中显示出来以便于用户分析不同地区用户的模式(关联性). 文章也设计了一个glyph, 将原本一个peak的多维信息结合在一起. 这个glyph下面和右边的线分别是时间线, 中间的渐变色块代表了分布, 一根横线的位置和粗细又编码了其他内容. 这个设计我个人很喜欢, 主要是简洁并且颜色美观.

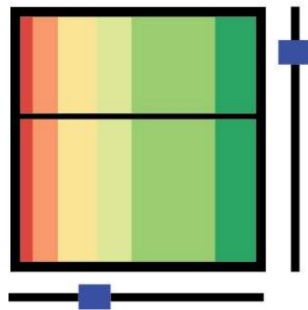


图 2: PeakVizor的glyph设计

- VAST14 *DecisionFlow: Visual Analytics for High-Dimensional Temporal Event Sequence Data* 看这篇文章是因为也是下面一篇文章的baseline. 尽管两篇文章的侧重点并不相同, 但是其描述的领域都是医疗上的数据. 本文相当于是从可视化的角度, 对一种相关性进行分析与判断. 首先进行Visual Query, 得到一些想要的数据集, 对于得到的数据, 还可以人为的交互进行添加一些milestones, 相当于是决策树中的节点. 当然也有针对此的删除操作. 决策时具体的信息可以通过交互展现, 相关性和显著性等统计信息, 显著的反馈给用户进行下一循环的探索. 整套流程很好的体现了可视数据挖掘的特性.
- IUI16 *Adaptive Contextualization: Combating Bias During High-Dimensional Visualization and Data Selection* 这个会议ACM IUI全称intelligent user interfaces, 其自我介绍为HCI和AI遇见的地方. 文章做的内容是高维数据的选择. 有点类似于以往有不少可视分析领域文章做的是特征选择, 而本文是在

数据选择上做文章。由于如今的多数数据维度都很高, 有时候我们在数据选择时候可能会有意无意的由于没有注意到不同特征内在的联系而使得选择完以后的数据和之前原本的数据集有bias。举个例子, 对于病人的数据, 我们如果选择时候只选经历紧急救助的, 那将会导致数据skew到那些急性病的人身上, 这样的选取就不是无偏的。为此, 需要度量每次过滤(选取)后的结果与之前的数据的差距, 这里用了Hellinger distance, 用于度量两个分布的距离(这可以联想到我们之前课程学的Earth Mover's Distance, 同样作为一种分布差异的度量方式)。此外在此基础上还定义了多multi-variate measure。从这里隐约可以看到那种类似于以前那些特征选取的套路。文中定义的Provenance model, 其实作为一个一步步过滤的过程, 类似于一个决策过程, 也就可以直接套用上面的DecisionFlow了。同样的本文的Glyph设计思路也不错, 尽管我觉得可能叫small multiple更适当些。它作为一种小overview其实包含了足够的信息帮助一步步的决策。

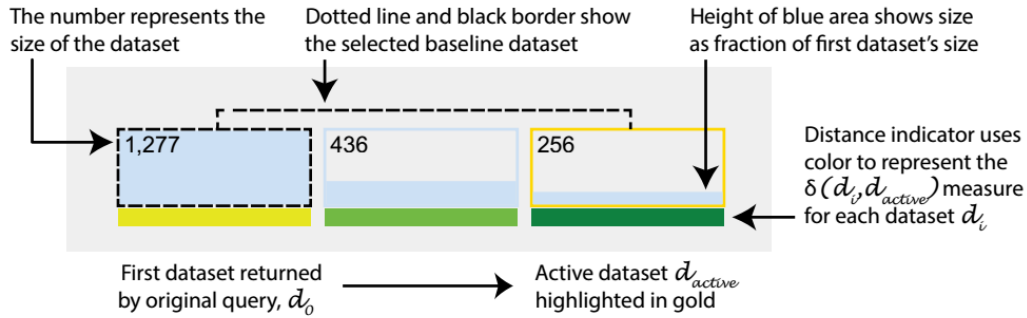


Figure 4. The contextualized breadcrumb view uses color-coded bars at the bottom of each glyph to encode the δ measures. Interaction capabilities revert to prior datasets or select a prior dataset for detailed comparison in the Balance Panel.

图 3: Adaptive Contextualization的glyph设计